Observations on CIA's Efforts in Developing Analytical Methodologies

**A Report of the DCI's
Science and Technology Advisory Panel**

SCIENCE AND TECHNOLOGY ADVISORY PANEL

OBSERVATIONS ON CIA'S EFFORTS IN
DEVELOPING ANALYTICAL METHODOLOGIES

1. Background

John Hicks, when he was DD/NFAC, requested that STAP
examine CIA's work on analytical methodologies. In response
to this request members of STAP met with representatives of
ORD, OSI, OWI, OSR, OPA, OGCR, OER, and ODP. As a result of
these discussions and an analysis of past and current CIA
efforts in this field, we make a number of general
observations. In addition, we identified four topics that we
discuss in greater detail:

   o Evaluation of Analytical Efforts
   o Analyst-User Interaction: Feedback
   o Role of Automation
   o Interdisciplinary Analysis

Moreover, an exemplar computer based system, SAFE
(Secure Analysts' File Environment), was examined to some
depth; it is the subject of a separate report.

2. General Observations

Analysts are individuals and each carries out his work
in his own way. The development of analytical methodologies
requires some understanding of the common features of how
analysts work. However, at present there is no common
language/framework for describing analysis. Further, there
are no commonly accepted measures or approaches to
characterize analytical efforts (See Section 3). What
assessment is carried out tends to be performance-oriented
rather than value-related. The systematic comparison of
analytical methodologies and validation of results is
difficult since different measures and approaches are used.
The development of analytical methodologies is hampered by
lack of coordination among some of the responsible offices
and lack of commonality of perception by the managers of the
analytic efforts.

A recurring theme in the discussions was the
statistical nature of intelligence. The analyst works from a
limited sample of relevant data and the analyst is sometimes

aware of the bias in the sample. Despite these views, quantitative statistical methodology is little used within CIA. In the political area in particular, analysts generally are aware of the quantitative methods of political science but not much use is made of them, in part because of the heavy premium placed on current intelligence analysis and production. An initial ORD effort is only now being made to put together a base of statistical data (e.g., census, statistical yearbooks,) which are essential for quantitative political analysis.

The support of statistical analyses and other quantitative methodologies by on-line computers is widespread in our culture, as in medical research, chemical industries, seismology, and so on. In the Intelligence Community, such tools have been available and utilized very unevenly -- but the integrative nature of intelligence means that the real savings can come only when all the files are computer accessible. The concept of SAFE is partly directed at that.

We found no consensus as to whether the development of analytical methodologies should be centralized with "off line" development of the methodologies (ORD) or considered part of the continuing work of the NFAC offices, such as OSR, where methodologies are developed in the course of problem solving. Either or a combination of both can be made to work, but only if there is a manager with the responsibility to see to it that development of appropriate methodologies does indeed take place. In fact, no one at present has that responsibility. We are seriously concerned that the present largely haphazard approach to analytical methodologies will be perpetuated unless management responsibilities and oversight are clearly defined and understood by the participants.

3. Evaluation of Quality of Analytical Efforts

We have found almost no cases where an evaluation procedure is customarily applied to finished intelligence output in a rational or well-understood way. Finished intelligence is the primary output of the Intelligence Community, yet there is no rational way of properly evaluating contributions to that output. Derivatively, therefore, methods of analysis and even the analysts themselves cannot be consistently and appropriately rated.

That further implies difficulties in the rational hiring and selection of analysts and their optimum training. It is essential, therefore, to exercise proper evaluation of the outputs and components of intelligence processing.

This is not to say that it is impossible to find outright errors in analysis -- like failing to make a diligent search of sources; but it means that such failures surface only when the main thrust of the work is wrong. A correct conclusion will excuse the worst errors in interpretation and inference, because they will probably never have been noticed. Efficient evaluation of quality needs to be applied to three kinds of objects:

1) The intelligence document; that is, the individual output from the intelligence analysts.
2) The intelligence analyst; that is, the individual practitioner of intelligence analysis.
3) The intelligence analysis; that is, the set of analytical procedures and resources.

Evaluation, of course, is not useful in itself; it serves several ends. Nearly always, an analyst at the Agency is dedicated, hard-working, and responsive. Rapid, supportable evaluation of a document as a matter of course leads to improvements not only in that document but in succeeding ones. Furthermore, feedback about a report's influence in high places is a very special kind of reward that ought to be utilized more often. At another level, good evaluation enables a manager to know how to control his resources, including his analysts, and to deploy them in an optimum way. The exact nature of the "optimum" is prescribed to the manager by his manager or superior.

We can, then, make a list of the roles played by the evaluation of quality; it should be emphasized that the point of knowing that the worth of something is, say, 3.2, is that then one knows that when another thing is worth 3.7, it is better -- that is, the differences count, not the absolute level.

1)   If we can evaluate analysts in their tasks, their managers can assign them so as to maximize their joint effectiveness; an analyst's contribution can be improved by

changing his tasking or work conditions in a rational way.

2)   If we can evaluate analysts, it should be possible to begin to make more sensible and sensitive selection of analysts.

3)   Once selected, the analysts can be given training that is more relevant and that can be shown to improve them and their performance.

4)   The support of other resources, both automation and otherwise, can be fine-tuned according to how performance is improved, not merely according to some interpretation of doctrine.

5)   Finally, the larger structure of the analytical process can be improved, and known to be improved, only if the evaluative techniques are sufficiently accurate, precise, and responsive to the national needs.

Case studies -- an example of retrospective evaluation -- are a valuable tool. They may be, and often are, misused, as in "whom can we blame for 1) Afghanistan, 2) El Salvador or, 3) Iran?" Properly used, a case study should, inter alia, assign credit to the parts of the analytical process; it should compare alternative or conflicting processes or parts of processes -- by accurately evaluating their contribution in the case at hand.

4. Analyst-User Interaction: Feedback

The interaction between the analyst and the user can be used to improve substantively the quality of the analysis. Concurrently, such interaction can lead the user to frame requirements in a way that the Intelligence Community can best respond. However, the analyst-user interaction is an extremely delicate one. A user may come to depend on the quick reaction judgments of a few close associates when these judgments do not reflect the total community input. The analyst may come to know the users' views so well as to tailor his analysis to support those views and gain the

plaudits of the policymaking community. Despite these
dangers, the inherent values of the interaction are great
enough to warrant study.

In the course of our study, we identified a few cases
where the user was dissatisfied with the product, but this
dissatisfaction did not get back either to the analyst or
his manager. Nor were the requirements examined to determine
whether the user had communicated his needs in such a way
that the community really knew what he wanted.

Many users either have been in the Intelligence
Community or know the community well. Some users have had
little or no experience in intelligence but have good
experienced intelligence officers assigned to their staffs.
Still others have neither experience nor continuing contact
with the Intelligence Community.

We conclude that the user-analyst interaction is
important, and current practices, successful and
unsuccessful, need to be examined. The examination should
include experienced and inexperienced users, both short-term
and longer term estimates, and short and long turnaround
times. Such a group of retrospective looks at user-analyst
interaction could lead to guidelines or suggestions both to
the user and the Intelligence Community of how best to
interact, remembering at all times that the users, analysts,
and managers are people with individual talents, styles, and
shortcomings.

5. The Role of Automation in Analytical Methodologies

There is little doubt that automation, including all
the various tools and powers of modern technology, can
vastly enhance the processing capabilities. This section
examines what ought to be expected of automation, what can
reasonably be expected, and what ought not to be expected of
automation.

By automation we mean in this context mostly the
application of modern computer technology: it brings
problems as well as powers, both technical and human. It
enables the analyst to do things in ways and scales and
times that could not otherwise be dreamt of. But the
analysts must also change what they do. In some ways they
must become more vulnerable. In the large, the net gain is

enormously positive, so that we endorse the current drive toward computerization of many of the analytical processes that now are being undertaken.

There are gains to be made and pitfalls to be avoided. We have already discussed the urgent need to find out what the analytical processes consist of; that is, what analysts do, how to tell whether they are doing it well, and how to tell whether a change is an improvement. That study is necessary as well in order to plan, design and control the computer resources assigned to analytical processes in a responsible way. Indeed, the use of computers online can itself help to gather data about analytical processes, once they have been properly matched to the task. A beginning taxonomy of functions can be made:

1. File Management - It is widely acknowledged that there is little consistency in handling intelligence files throughout the agency; some analysts maintain "shoe boxes," some maintain computerized files. And many analysts use both those and others. It must be said emphatically that those techniques are not necessarily inefficient or poor, with respect to the context in which they are used. For most, certainly, they have been finely tuned to their contexts, the capabilities of the individual analysts, and the current demands they face. But some of the disadvantages can be remedied with sensitive use of computerized files:

   a. Standards and protocols of retrieval are totally personal and ad hoc. This means that every file takes individual work in setting up and updating; indeed it is well known that that is one of the most essential and tedious jobs that analysts do. If a new kind of data shows up, the analyst himself must make intellectual decisions about how to handle it. The merging of files is nearly impossible; merging the files from two analysts is usually totally impossible.

   b. As the circumstances of the task change, there is no easy way to modify the behavior of the file; for example, a file built around the retrieval according to certain attributes must be restructured manually and laboriously by the analyst himself, if the relevant

attributes change with the political
situation.

c. Sharing of files among analysts is rarely
feasible.

2. Retrieval - This is a task acknowledged to have a
large element of the mechanical. Done manually, it is very
prone to human error. Analysts remember things in their
heads according to attributes almost certainly not the same
as the ones they built into the file system at its
beginning. Furthermore, with large files, it can take a lot
of calendar time, and there is no way to speed it up*.

3. Aggregating, collating, sorting, disambiguating -
This includes the more or less mechanical matching and
averaging of items in a file; like adding up dollars to
produce a total for agricultural output.

4. Communication - Information in an agency's files
belongs to the agency and to the entire community. Far more
tedious even than retrieving from one's own files is
retrieving from distant ones in another agency.

5. Distribution - Proper distribution of collected
information must be regarded as an essential part of the
analytical process. Subjective judgment is not necessarily
poor at distribution, but much of it can be done routinely
without the high-powered expensive judgment of trained
analysts.

6. Presentation Means - The ultimate evaluation of the
agency derives from its output, which at present nearly
always means paper documents. One of the biggest delays of
production of documents comes from printing and reviewing.
For some documents, the calendar time spent on that is more
than the analyst spends in producing the original.

7. Comprehension, Assessment and Inference - This
category includes the most human functions of analysis. It

---

*Calendar time means real duration as opposed to computer
time (cpu cycles) or manhours.

is the high level effort that can be neither delegated nor automated, though it can be assisted.

The use of computers can obviously assist some of those functions more than it can in others. Computer programs cannot now "comprehend" anything, let alone intelligence information, in any human sense; while they are widely used in collating, etc. The capabilities currently proposed for SAFE go part of the distance towards those functions, and are the subject of another report.

Computers are far more effective when they are embraced by the users than when users are coerced. It should be realized that even with a large and expensive computer system, the users, that is, the analysts, are by far the most expensive part of the whole system. This means that the imposition of rules and regulations about the acquisition of equipment, its utilization, how terminals are to be assigned, and so on, must be continually sensitive to a rational evaluation of productivity of the whole system, both analysts and computer systems. We do not know now what the best way will be to distribute and assign computing power, and we cannot know it until we try. Furthermore, computing will not merely change the costs of doing tasks, but will also change the kinds of tasks attempted, and the way they are done. The most effective way to set standards is to reward the improved productivity that comes from following them, and that is the primary way management should exercise its powers, not by requiring certain kinds of equipment to be used at certain moments of operation.

Analytical methodologies should be exploring current techniques of processing so as to help in the transition to a more efficient utilization of processing where it is profitable. Such studies can break loose from current attitudes and practices, for it is well known to be difficult for an analyst alone to stand back from his environment of continuing crises and pressures and to examine the broader aspects of change and computerization.

It is tempting to run through the list of functions above, showing how computers can, or cannot, help in each item; but this is probably not the place to do that, and the agency has, in many cases, already done so. What should be emphasized is that the traditional breakdown of functions will be altered by computers and often in significant ways.

For example, communicating computers that can transfer files easily can enable remote collaboration in which analysts do not have even to see each other in order to collaborate interactively in intelligence analysis--indeed, they might be in different agencies. But automatic inferencing is not one of the functions that will soon be automated, although it should certainly be studied.

Similarly, evaluation is far easier in a computer environment, just because the extra load is negligible for the computer to store the actions and queries that an analyst performs, and to group them according to category. The action and queries can be a powerful tool in analyzing the procedures used by an analyst. If the analyst is working interactively, which we presume ought to be the usual mode, then merely storing each query/request, together with the exact time it was issued, will provide adequate data. The study of those data could well serve as the keystone to a truly dynamic effort to develop new analytical methodologies.

The role of management should be one of leadership not prescription. It should not, for example, prescribe the use of this or that computer language; rather, it should point out the advantages of being able to share programs and data with other users' files. Part of the reward structure for compatibility with other units in the agency could well be organizational or budgetary in nature--equipment or software that enhances compatibility and collaboration could be supported by higher level management.

A key theme is powerful communication, which both supports and is supported by computerization. The use of message systems agency-wide (community-wide?) can add new dimensions to the possibility of collaboration, to say nothing of distribution and presentation. Furthermore, the structure of the message system can go far toward assuring the compatibility of different systems: the difference is between requiring the use of a particular machine and requiring that any system used must be able to participate in a message system.

Fundamentally, automation and use of computers will not produce magic results, although they can produce very remarkable ones. One topic we have not touched on here is the integration of the information flow from collection to

processing, so that the format in which information is collected and distributed should more nearly match those needed for efficient processing. The advent of very cheap micro processors means that reformatting or translating data, even in real time is neither expensive nor time-consuming, over a broad range of data types and rates. This will profoundly influence the direction followed by future integrated SIGINT systems for collection and processing.

6. Interdisciplinary Analysis

STATINTL

STATINTL

CIA/ORD has developed new methodological approaches to multidisciplinary problems in two highly important areas; Soviet Grain Production ███████████████ and Soviet Oil Reserves ██████████████ Both cases involved questions that clearly required a mix of expertise in order to obtain meaningful answers. Economists, agricultural biologists, or geophysicists by themselves could not have adequately dealt with either problem.
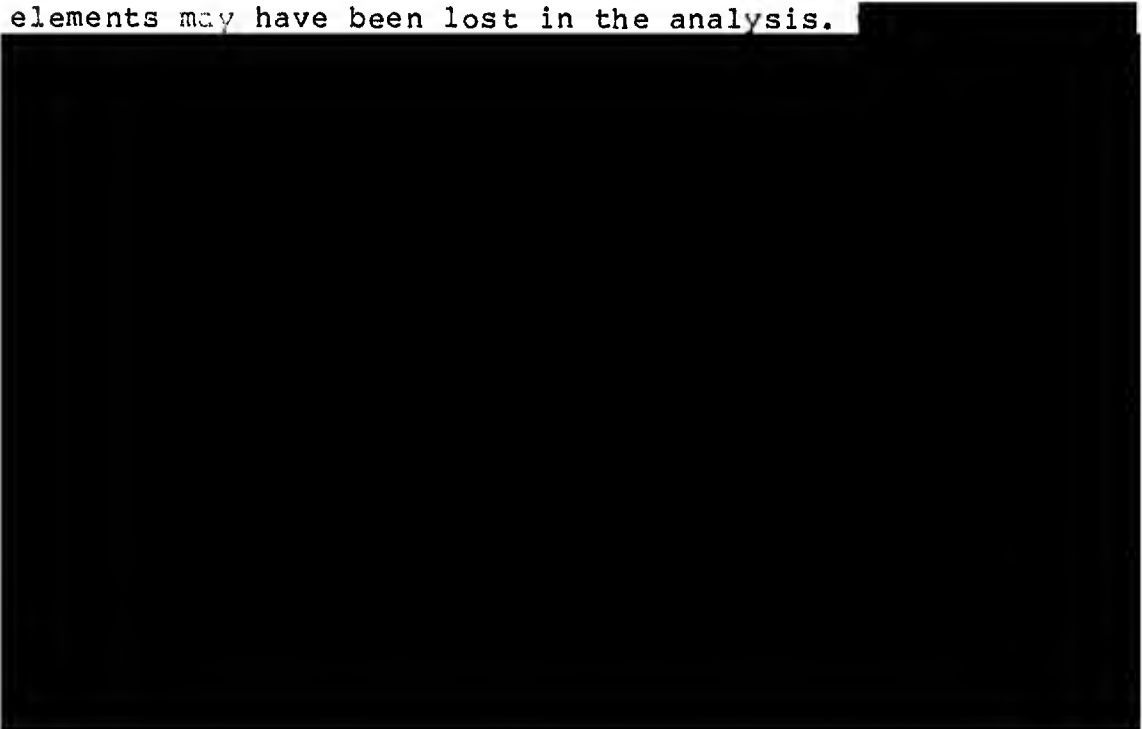
An important element of both examples is that the character of the question drove the method of the analysis. The experience in OSR and the close working relations between OER, OSI (now OSWR) and OSR illustrate the point that the nature of the question influences the way the problem is approached.

STATINTL

STATINTL

The "style" of OSR is to phrase or rephrase the question in such a way that the more general aspects of the problem are highlighted. In OSR, working as a member of multidisciplinary team does not, it appears, negatively influence the rewards; in fact the opposite seems to be true. The OSR structure and management have historically encouraged not only the recruitment of analysts with a variety of backgrounds, but also the contribution of the analyst to group efforts, and his cooperation with other offices within NFAC.

Problems of military importance have often been viewed in relatively narrow terms and as a result important

STATINTL

elements may have been lost in the analysis.

        Non-military questions involving issues such as natural
resources are newer to the Intelligence Community and since
they are less well defined, have lent themselves to
multidisplinary analysis. As these topics develop, so will
the specialist, and there is a danger that the important
interactions may be lost in increasing sophistication of the
analysis. This is a danger that Intelligence Community
managers must be aware of and guard against.

        We conclude that multidisciplinary efforts are natural,
often useful, and sometimes essential. These efforts can be
encouraged in a variety of ways:

        o   Questions should be phrased by the user or
                rephrased by the community in such a way as
                to highlight the interconnections among the
                issues related to the initial question. Top
                level NFAC management, including NIC, must
                play a key role in properly structuring the
                questions.

STATINTL

o   Managers should structure their operation and
      in particular the reward system so that
      multidisciplinary and interoffice analytical
      efforts receive appropriate recognition.

o   The offices involved (OER, OGCR, ORD, OSWR)
      should carry out a retrospective analysis of
      ████████████████████ both in terms of how "good"
      the results were but also what were the key
      elements in the management of the
      multidisciplinary activity.